

## LA-UR-21-26671

Approved for public release; distribution is unlimited.

Title: LUNA Condition Based Monitoring Update: Using Minimum of Mahalanobis Distances for Multi-Class Classification

Author(s): Green, Andre Walter

Intended for: Progress report to sponsor

Issued: 2021-07-13

---

**Disclaimer:**

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

# **LUNA Condition Based Monitoring Update:**

## Using Minimum of Mahalanobis Distances for Multi-Class Classification

Presented 7/13/2021

# Multi-Mahalanobis Classifier | Concept

The value of  $r$  in the equation

$$r^2 = (x - m_x)' C_x^{-1} (x - m_x) \quad (4)$$

is called the **Mahalanobis distance** from the feature vector  $x$  to the mean vector  $m_x$ , where  $C_x$  is the covariance matrix of  $x$ ; it can be demonstrated that surfaces in which  $r$  is constant are ellipsoidal with center in  $m_x$ . The Mahalanobis distance can be used in a minimum distance classifier as following (Figure 1): be  $m_1, m_2, \dots, m_n$  the centroids of the  $n$  classes, and be  $C_1, C_2, \dots, C_n$  the corresponding covariance matrixes. A feature vector  $x$  is classified by measuring the Mahalanobis distance from  $x$  to each of the centroids, and by attributing  $x$  to the class in which the Mahalanobis distance is minimum [5].

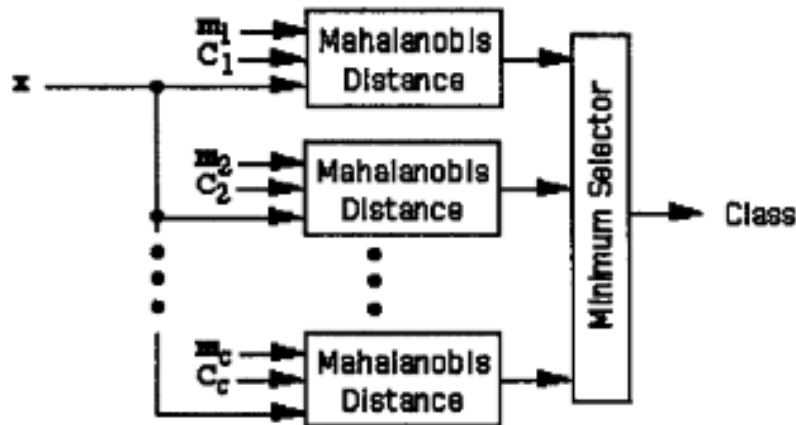


Figure 1. Mahalanobis distance in a minimum distance classifier.

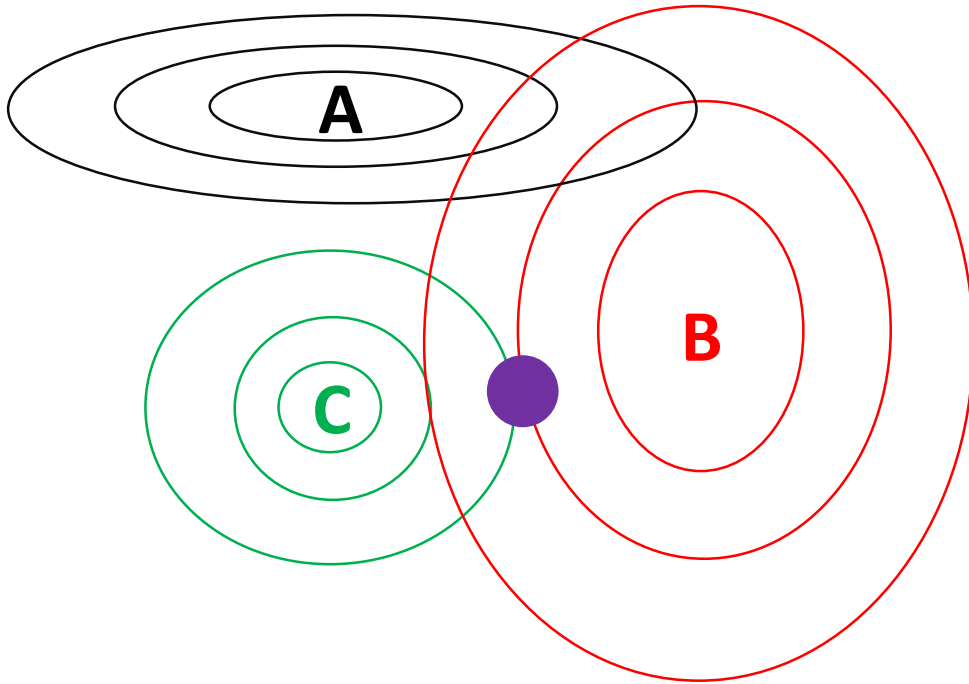
“After an initial learning period, the algorithm extracts the four features from every new QRS\* complex and calculates the Mahalanobis distance between its feature set and the centroids of all existing classes to determine the class in which the new QRS belongs to.

If a predefined distance is surpassed, a new class is created.”

<https://ieeexplore.ieee.org/abstract/document/1166742>

*\* The ‘QRS Complex’ is a combination of the Q, R, and S wave types on an electrocardiogram. These correspond to the depolarization of the left and right ventricles of the heart and the contraction of some heart muscles.*

# Multi-Mahalanobis Classifier | Concept



*Illustration of multiple Mahalanobis ellipses. Each ring represents a different Mahalanobis distance to the center of a given type (indicated by a letter).*

The purple point would be classified as type B, in the example above. Although it may appear equally close to C, it's much closer relative to B given the scaling applied by B (i.e. in terms of the rings the point is closer to B – it's right on the edge of its second ring).

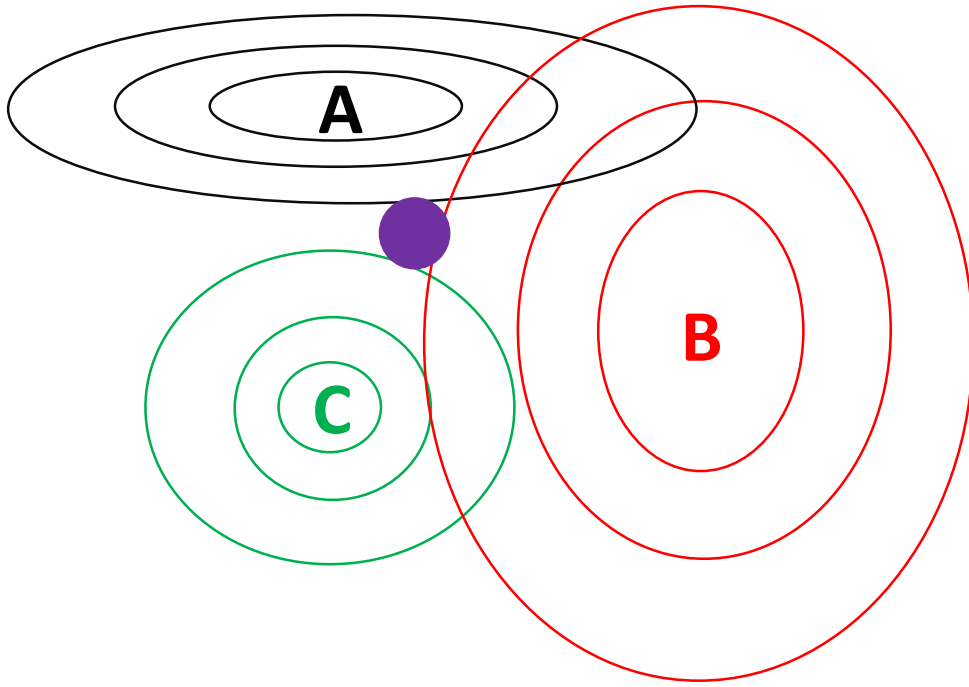
For a point P, the vector of Mahalanobis distances is defined as:

$$\langle \text{MHL\_A}(P), \text{MHL\_B}(P), \text{MHL\_C}(P) \rangle$$

Where  $\text{MHL\_N}(P)$  is the Mahalanobis distance from the point P to the mean of the distribution N, with the axes scaled by N's covariance matrix.

The distribution (i.e. damage type) corresponding to the lowest distance is selected as the damage type for the point P.

# Multi-Mahalanobis Classifier | Concept



*Illustration of multiple Mahalanobis ellipses. Each ring represents a different Mahalanobis distance to the center of a given type (indicated by a letter).*

The purple point could be classified as any of A, B, or C. It's reasonably close to all of them, so its minimum distance is not high: however, because the distances are so similar (distance is interpreted as the probability of belonging to a class), the entropy should be high.

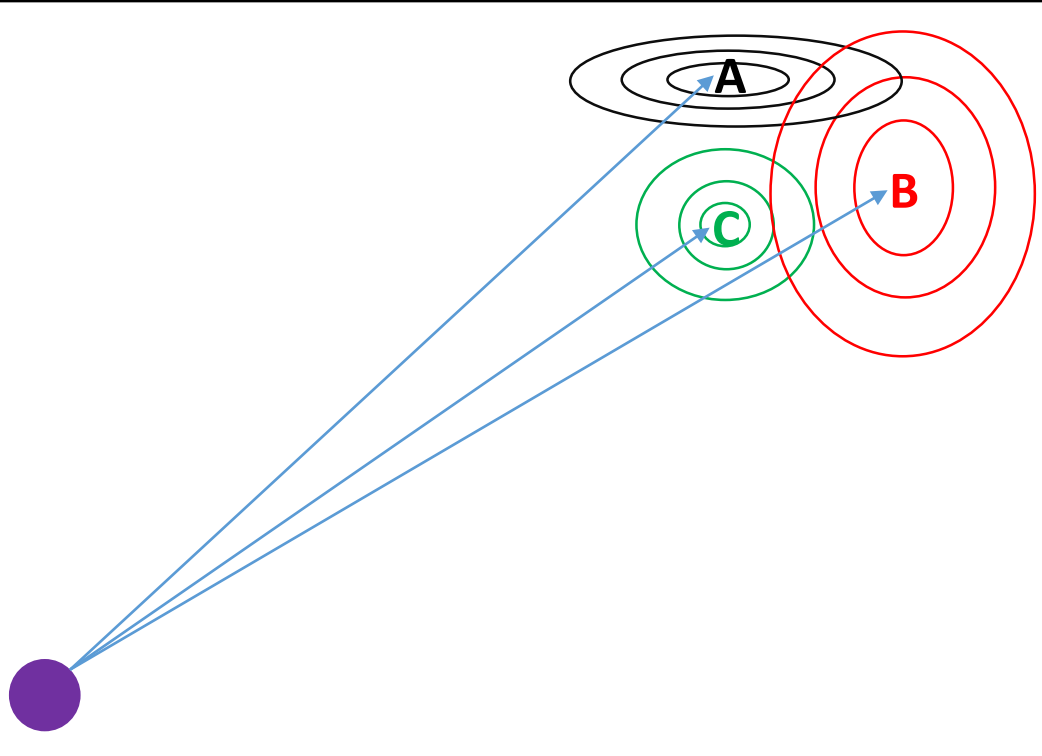
For a point P, the vector of Mahalanobis distances is defined as:

$$\langle \text{MHL\_A}(P), \text{MHL\_B}(P), \text{MHL\_C}(P) \rangle$$

Where  $\text{MHL\_N}(P)$  is the Mahalanobis distance from the point P to the mean of the distribution N, with the axes scaled by N's covariance matrix.

The distribution (i.e. damage type) corresponding to the lowest distance is selected as the damage type for the point P.

# Multi-Mahalanobis Classifier | Concept



*Illustration of multiple Mahalanobis ellipses. Each ring represents a different Mahalanobis distance to the center of a given type (indicated by a letter).*

The purple point is far from everyone, so even its minimum distance is high. Likewise, If the distances are all high for all types, the entropy of the distance vector becomes high.

For a point P, the vector of Mahalanobis distances is defined as:

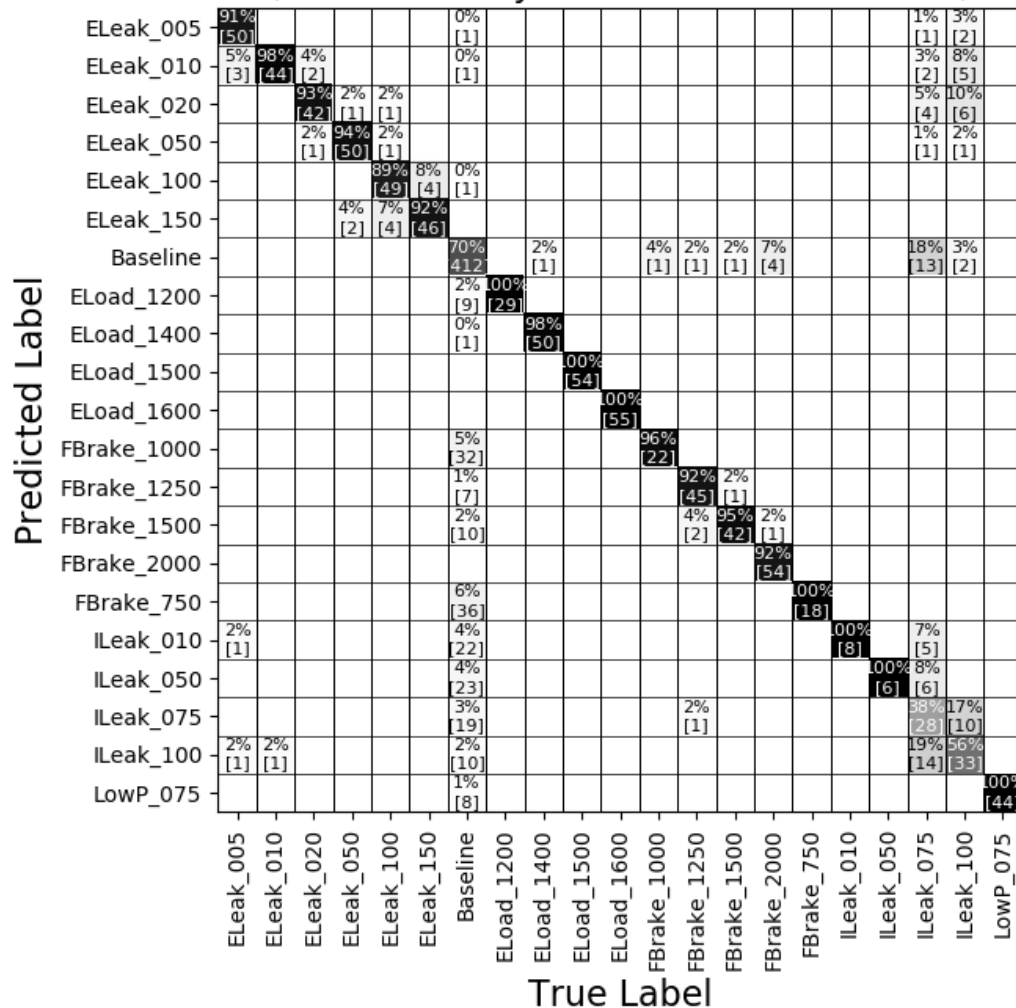
$$\langle \text{MHL\_A}(P), \text{MHL\_B}(P), \text{MHL\_C}(P) \rangle$$

Where  $\text{MHL\_N}(P)$  is the Mahalanobis distance from the point P to the mean of the distribution N, with the axes scaled by N's covariance matrix.

The distribution (i.e. damage type) corresponding to the lowest distance is selected as the damage type for the point P.

# Multi-Mahalanobis Classifier | Results

Confusion Matrix  
(Normalized by the sum of the columns)



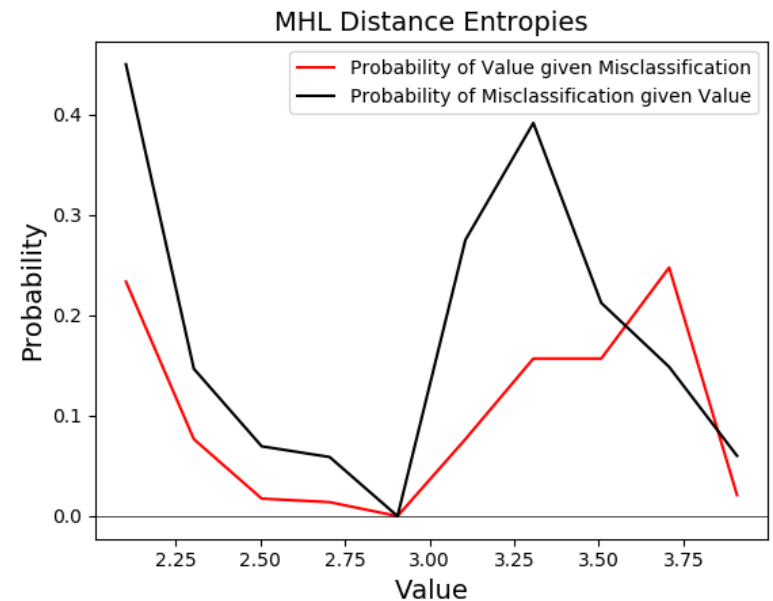
*ILeak 075 and 100 tend to get confused with each other most often, but also get mixed up in the baseline cases.*

## Using 9-Fold Cross Validation (Including ILeak cases)

Acc. 83.49, Acc. 82.82, Acc. 78.83  
Acc. 80.37, Acc. 82.21, Acc. 79.45  
Acc. 80.67, Acc. 74.85, Acc. 76.69

## Mean of 79.93%.

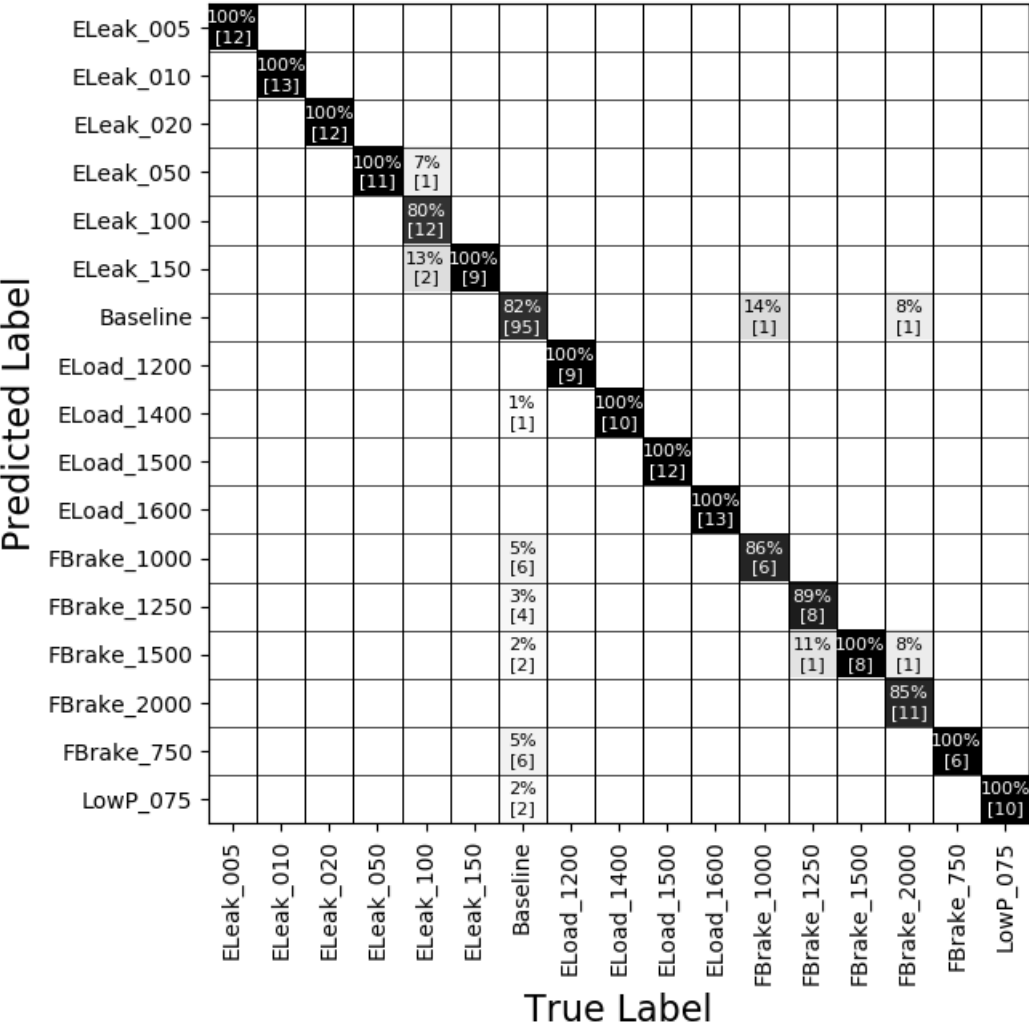
A Mahalanobis transform was made for each severity class, and each point was classified as belonging to the class whose transform brought it closest to the origin.



*If a sample has high entropy, it tends to get misclassified, and if it's misclassified, it tends to be at a high entropy. This is good: it's implying that the misclassified samples are the ones the network is already unsure about (uncertainty is expressed by high entropy across the Mahalanobis distances).*

# Multi-Mahalanobis Classifier | Results

Confusion Matrix  
(Normalized by the sum of the columns)

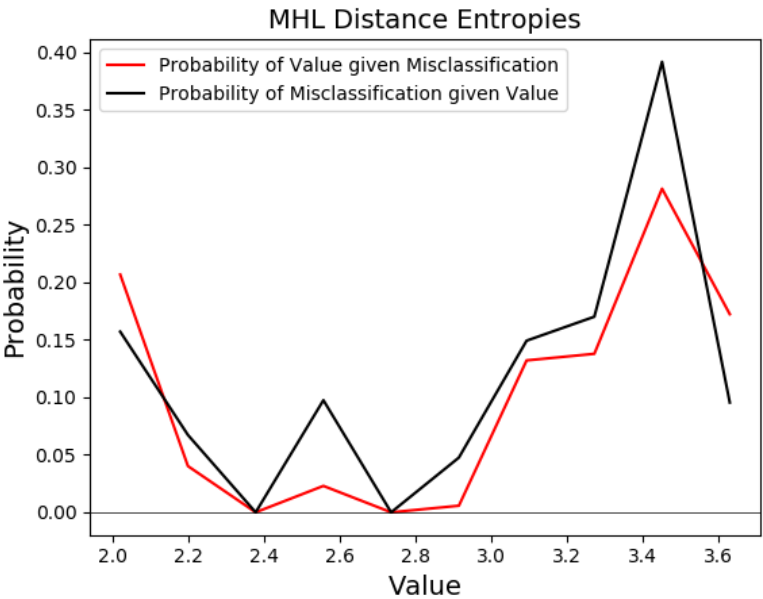


## Using 9-Fold Cross Validation (Not including ILeak cases)

Acc. 90.18 Acc. 91.23 Acc. 90.18  
Acc. 89.47 Acc. 92.63 Acc. 88.03  
Acc. 89.08 Acc. 91.2 Acc. 90.14

### Mean of 90.23%.

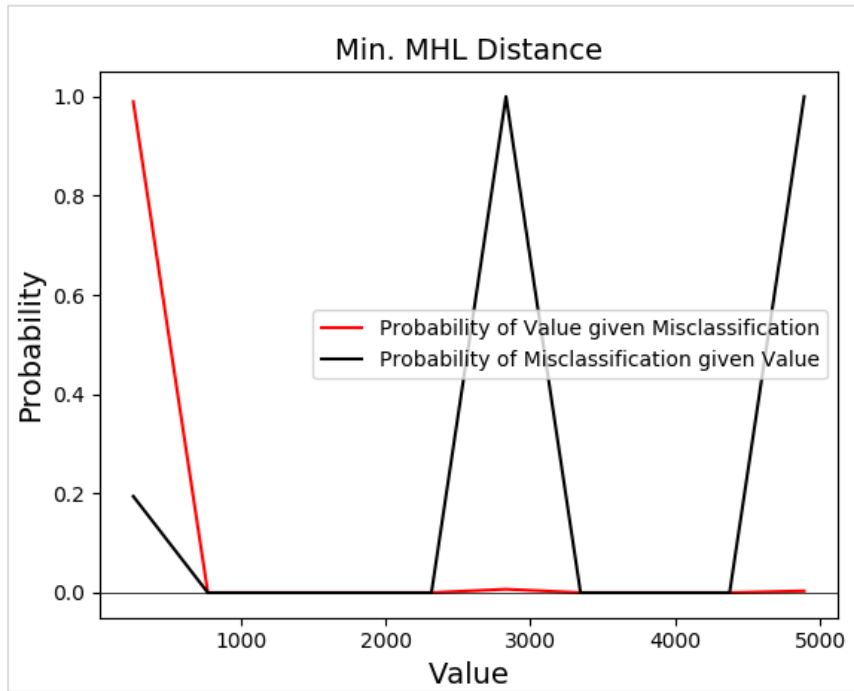
A Mahalanobis transform was made for each severity class, and each point was classified as belonging to the class whose transform brought it closest to the origin.



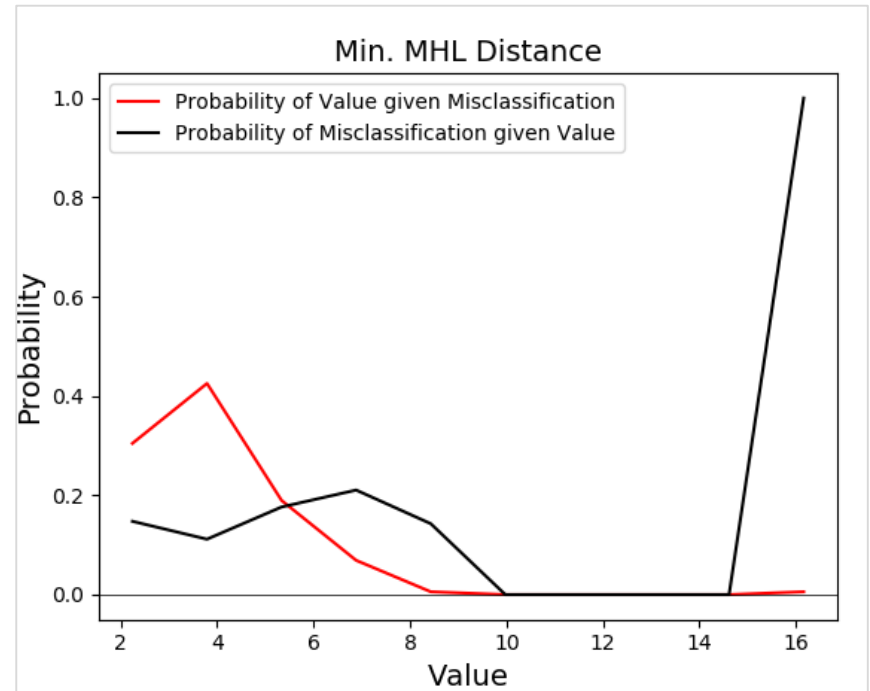
If a sample has high entropy, it tends to get misclassified, and if it's misclassified, it tends to be at a high entropy. This is good: it's implying that the misclassified samples are the ones the network is already unsure about (uncertainty is expressed by high entropy across the Mahalanobis distances).

# Multi-Mahalanobis Classifier | Results

## Including ILeak



## Excluding ILeak

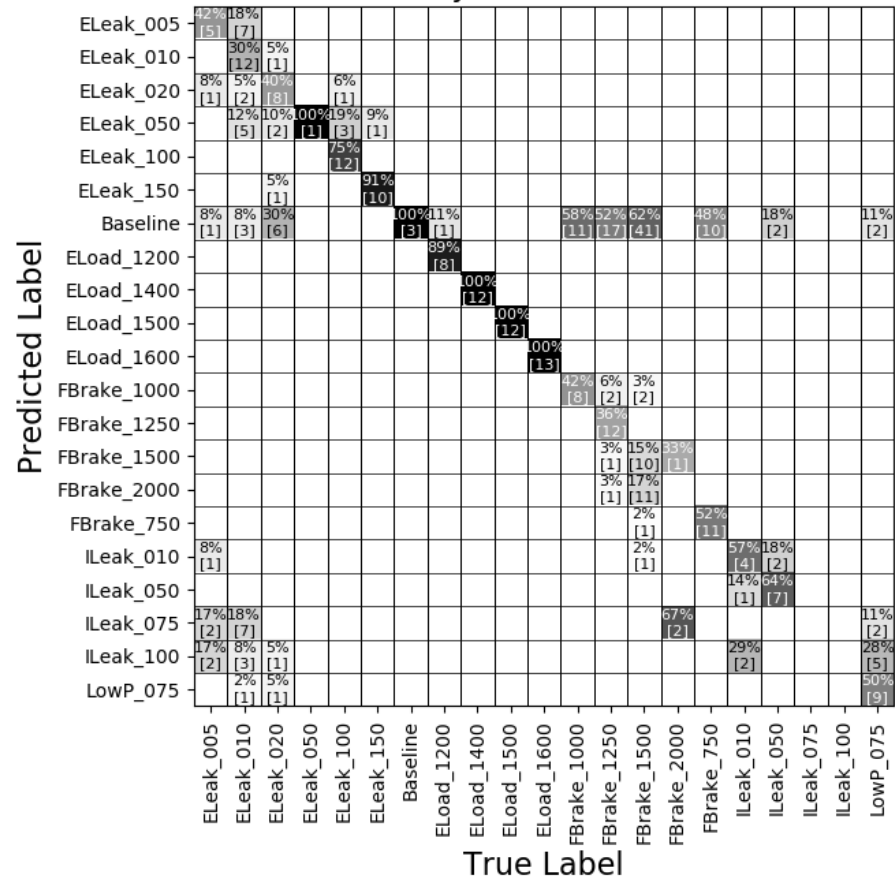


The x-axis goes way higher on the graph including ILeak, indicating there are at least two cases that are genuinely anomalous (as per the QRS paper referenced previously). These points are misclassified

For both graphs, the red line is initially above the black, but later the black is above the red: i.e. most misclassified samples are on the lower end (tending to have low minimum distances), but those that are on the higher end (long minimum distances) have a greater chance of being misclassified.

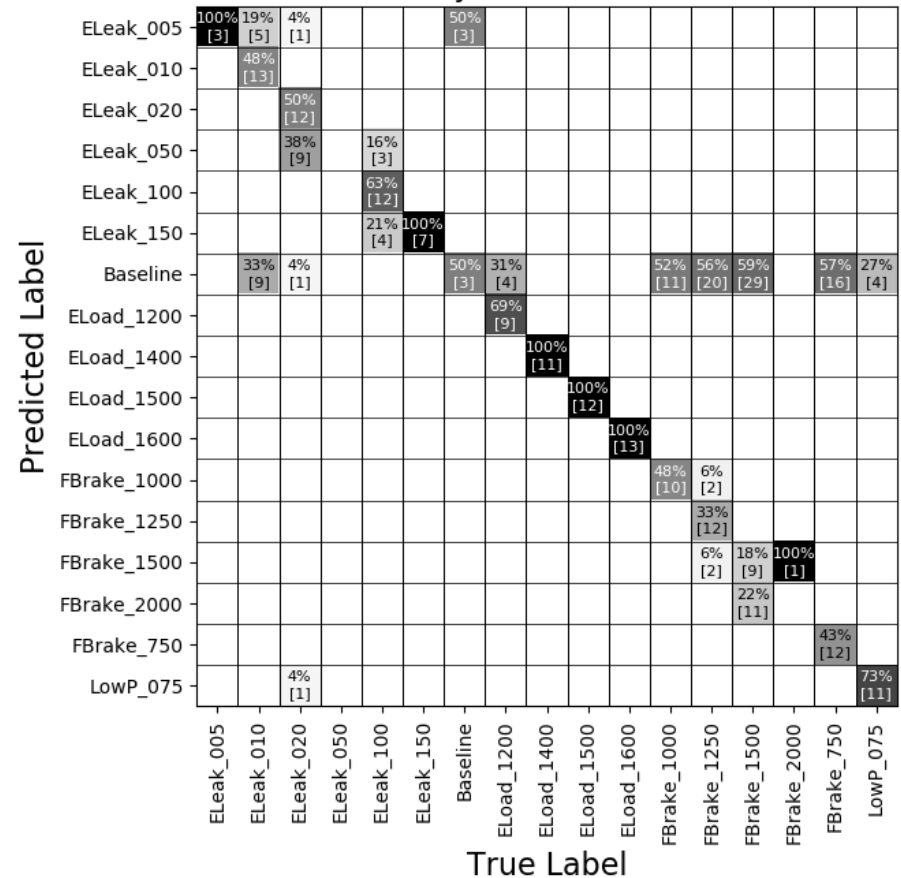
# Multi-Mahalanobis Classifier | MinCovDet (Defaults)

Confusion Matrix  
(Normalized by the sum of the columns)



With ILeak, MinCovDet

Confusion Matrix  
(Normalized by the sum of the columns)



No ILeak, MinCovDet

Using robust covariance (MinCovDet) performs worse than regular covariance: this may be because the data is not distributed uni-modally. From the sklearn documentation, “[MinCovDet] is not meant to be used with multi-modal data (the algorithm used to fit a MinCovDet object is likely to fail in such a case).”

<https://scikit-learn.org/stable/modules/generated/sklearn.covariance.MinCovDet.html>

# Multi-Mahalanobis Classifier | Papers

“Closed-Form Training of Mahalanobis Distance for Supervised Clustering”

[http://www.cs.toronto.edu/~law/publications/CVPR/2016/mlca\\_cvpr\\_2016.pdf](http://www.cs.toronto.edu/~law/publications/CVPR/2016/mlca_cvpr_2016.pdf)

“Adversarial Mahalanobis Distance-Based Attentive Song Recommender”

<https://arxiv.org/pdf/1906.03450.pdf>

Learning a Mahalanobis Distance Metric for Data Clustering & Classification

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.703.4133&rep=rep1&type=pdf>

An Extreme Learning Machine Method for multi-classification with Mahalanobis Distance

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8469606>

**“Exploiting Multiple Mahalanobis Distance Metrics to Screen Outliers” [Manufacturing paper which talks about outlier removal]**

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6227532>

“Multivariate Characterization of Morphological Traits in Burkina Faso Sheep”

<https://reader.elsevier.com/reader/sd/pii/S0921448808002034?token=0F3BEB1CC3C6C018752043D71792B61163F5CAAB9D53FF5621A7E41243D73A710A1D91D3683FFE76C099DDE65AD9A722&originRegion=us-east-1&originCreation=20210712223622>

**“A real-time QRS Complex Classification Method using Mahalanobis Distance” [EEG Paper which uses the same technique as these slides]**

<https://ieeexplore.ieee.org/abstract/document/1166742/>

“Minimum Mahalanobis Distance Functions and Lithic Source Characterization by Multi-Element Analysis”

[Link is really long, but the paper's by Leach & Manly]

“A weighted Minimum Distance Classifier for Pattern Recognition”

<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=332440>

“Mahalanobis Distance-Based Classifiers are Able to Recognize EEG Patterns by Using a Few EEG Electrodes”

<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1019019>

“Max-Mahalanobis Linear Discriminant Analysis Networks”

<http://proceedings.mlr.press/v80/pang18a.html>

# Multi-Mahalanobis Classifier | Web

“After an initial learning period, the algorithm extracts the four features from every new QRS complex and calculates the Mahalanobis distance between its feature set and the centroids of all existing classes to determine the class in which the new QRS belongs to. If a predefined distance is surpassed, a new class is created.”

<https://ieeexplore.ieee.org/abstract/document/1166742>

Multi-Mahalanobis seems kind of like a fancy form of LDA.

LDA assumes equal covariance matrices – this multi-Mahalanobis distance method does not. Here’s a stack exchange question about it:

<https://stats.stackexchange.com/questions/387175/is-lda-just-selecting-the-minimum-mahalanobis-distance>

“In order to use the Mahalanobis distance to classify a test point, ... one computes the Mahalanobis distance to each class, and classifies the test point as belonging to that class for which the Mahalanobis distance is minimal.”

[https://en.wikipedia.org/wiki/Mahalanobis\\_distance](https://en.wikipedia.org/wiki/Mahalanobis_distance)